

IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE

Attorney Docket No.: Google-5/CON1 (GP-016-01-US)

Applicants: William PUGH, Monika H. HENZINGER

Serial No.: Not yet assigned

Filing Date: Herewith

Title: DETECTING DUPLICATE AND NEAR-DUPLICATE FILES

Examiner: Not yet assigned

Group Art Unit: Not yet assigned

Mail Stop Patent Application
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

S I R:

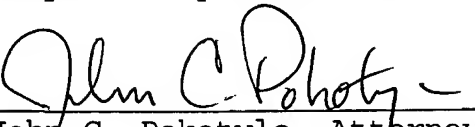
Information Disclosure Statement Transmittal

The applicants respectfully request that the references listed on the attached PTO/SB/08A be considered in the examination of the above-identified application. A copy of each of these references is enclosed. Each reference was cited in U.S. Patent Application Serial No. 09/768,947, of which the above-identified application is a continuation.

The applicants reserve the right to establish that any of the references listed on the attached PTO/SB/08A are not prior art to the above-captioned application.

Respectfully submitted,

Dated: June 27, 2003



John C. Pokotylo, Attorney
Reg. No. 36,242
Customer No. 26479
(732) 335-1222

STRAUB & POKOTYLO
1 Bethany Road
Suite 83, Bldg. 6
Hazlet, New Jersey 07730

*****EXPRESS MAIL CERTIFICATION*****

"Express Mail" mailing label number: EL780162162US
Date of deposit: June 27, 2003

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.


John C. Pokotylo

36,242
Reg. No.

(modified PTO/SB/08A)

| | | | | | |
|---|----------|----|----------|---|--|
| U.S. Department of Commerce Patent and Trademark Office INFORMATION DISCLOSURE STATEMENT BY APPLICANT <i>(use as many sheets as necessary)</i> | | | | <u>Complete if Known</u> Application Number: Not yet assigned Filing Date: Herewith First Named Inventor: William PUGH Group Art Unit: Not yet assigned Examiner Name: Not yet assigned | |
| Sheet | 1 | of | 2 | Attorney Docket No.: Google-5/CON1 (GP-016-01-US) | |

U.S. PATENT DOCUMENTS

[illegible][illegible]

| | | | |
|-----------------------|--|--------------------|--|
| Examiner Signature | | Date Considered | |
|-----------------------|--|--------------------|--|

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Unique citation designation number. 2 See attached kinds of U.S. Patent Documents. 3 Enter Office that Issued the document, by the two-letter code (WIPO Standard ST.3). 4 For Japanese patent documents, the Indication of the year of the reign of the Emperor must precede the serial number of the patent document. 5 Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16, if possible. 6 Applicant is to place a check mark here if English language translation is attached.

(modified PTO/SB/08A)

| | | | |
|--|---|--------------------------|------------------------------|
| U.S. Department of Commerce Patent and Trademark Office | | <i>Complete if Known</i> | |
| INFORMATION DISCLOSURE STATEMENT BY APPLICANT <i>(use as many sheets as necessary)</i> | | Application Number: | Not yet assigned |
| | | Filing Date: | Herewith |
| | | First Named Inventor: | William PUGH |
| | | Group Art Unit: | Not yet assigned |
| | | Examiner Name: | Not yet assigned |
| Sheet | 2 | of | 2 |
| | | Attorney Docket No.: | Google-5/CON1 (GP-016-01-US) |

| OTHER REFERENCES - NON-PATENT LITERATURE DOCUMENTS | | | |
|--|-----------------------|---|----------------|
| Examiner Initials* | Cite No. ¹ | Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume, issue number(s), publisher, country, where published, source | T ² |
| | AF | Andrei Z. Broder, Steven C. Glassman, Mark S. Manasse, Geoffrey Zweig, "Syntactic Clustering of the Web", <u>Proc. 6th Int'l. World Wide Web Conf.</u> , pp. 391-404 (April 1997) and <u>SRC Technical Note 1997-015</u> (July 25, 1997) (Downloaded from http://gatekeeper.dec.com/pub/DEC/SRC/technical-notes/SRC-1997-015-html/ on November 27, 2000, pp. 1-14.) | |
| | AG | Andrei Z. Broder, "On the resemblance and containment of documents", <u>Proc. of Compression and Complexity of Sequences 1997</u> , pp. 21-29 (IEEE Computer Society) pp. 1-9. | |
| | AH | Sergey Brin, James Davis, Hector Garcia-Molina, "Copy Detection Mechanisms for Digital Documents", <u>Proc. of ACM SIGMOD Annual Conf.</u> (San Jose, 1995) (Downloaded from http://www-db.stanford.edu/~sergey/copy.html on November 27, 2000, pp. 1-21.) | |
| | AI | Andrei Z. Broder, "Some applications of Rabin's fingerprinting method", R. Capocelli, A. DeSantis, U. Vaccaro, Eds; <u>Sequences II: Methods in Communications, Security, and Computer Science</u> , pp. 143-152 (Springer-Verlag, 1993). | |
| | AJ | Min Fang, Narayanan Shivakumar, Hector Garcia-Molina, Rajeev Motwani, Jeffrey D. Ullman, "Computing Iceberg Queries Efficiently", <u>Proc. 24th Int'l. Conf. on Very Large Databases</u> , pp. 299-310 (1998). | |
| | | | |
| | | | |
| | | | |
| | | | |

| | | | |
|--------------------|--|-----------------|--|
| Examiner Signature | | Date Considered | |
|--------------------|--|-----------------|--|

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Unique citation designation number. 2 Applicant is to place a check mark here if English language translation is attached.